Responses to the Community comment

We would like to thank to the reviewer for giving constructive criticisms, which are very helpful in improving the quality of the manuscript. We have made the revision based on the critical comments and suggestions of the referees. The referee's comments are reproduced (*black*) along with our replies (blue) and changes made to the text (red) in the revised manuscript. All the authors have read the revised manuscript and agreed with submission in its revised form.

Anonymous Referee #1

Specific Comments

Comment NO.1: Model Setup and Restart Timings

Has the accuracy difference in meteorological modeling used in WRF-Chem been considered for 24h, 48h, and 72h interval setups in with DA and without DA experiments?

Response: Yes. The experiments of Severe-FR-24h, Severe-FR-48h, Severe-FR-72h ingest the meteorological fields from the National Centers for Environmental Prediction Final (FNL) analysis data as the initial fields every 24, 48, and 72 hours respectively, and the corresponding DA experiments are conducted. The purpose of those experiments are to consider the accuracy difference in meteorological modeling used in WRF-Chem. The spin up on January 16 is for the meteorological fields.

Comment NO.2: Analysis vs Forecast Concept

Does the paper focus on reproduction rather than prediction? Clarify this aspect throughout the text. Does the 4D-LETKF approach via the equation at (line 152) include all the spatiotemporal information for xa or xb within the specified time-window? For example, if the 48h time-window assimilates 48 hours of 100x100x40 grid data, does it mean the results reflect only the analysis field? If forecasts are presented, the sustained impact of initial conditions should also be discussed.

Response: The paper focuses on reproduction and it has been clarified throughout the manuscript. The 4D-LETKF approach via the equation include all the spatiotemporal information within the window. The results reflect only the analysis field.

Comment NO.3: Ensemble Spread (Figure 9)

The ensemble spread represented by the standard deviation in the analysis results (third and fourth row) significantly decreases after assimilation. This is expected in ensemble-based data assimilation. I understood that the $\bar{x}^a + X^a$ is used for the next cycle. Is the inflation technique unnecessary, relying solely on emission perturbations for sufficient spread?

Response: Inflation technique is a classical method and benefit to make the information spread in analysis. If there is no Inflation technique, the assimilation result may overconfidence to model. Emission perturbations also facilitate the information spread, but it may not mean that the inflation technique is unnecessary.

Comment NO.4: Regional Labels

The paper mentions many Chinese regions. Figures 6, 7, and 9 zoom into BTH, but BTH labels should be added. Additionally, Figure 10 extensively discusses Shijiazhuang, which is not clearly located.

Response: Figures 6, 7and 9 introduce the BTH region in the revised manuscript. I also point out the exactly location of observation in Shijiazhuang in manuscript and Figure S5.

Comment NO.5: Terminology for Ensemble Spread

Section 3.3 uses various terms to describe the spread, such as divergence, convergence, dispersion, and high standard deviation. Please standardize terminology to avoid confusion and replace "divergence" with "ensemble spread". From the line 560, the paper mentions “filter convergence” instead of “filter divergence”, which is well known terminology. As you know, the filter divergence occurs when ensemble spread becomes too small, preventing observations from impacting the model.

Response: I have standardized terminology to avoid confusion and replace "divergence" with "ensemble spread". I have solved the relevant problem, thank you.

Technical corrections

Comment NO.6: (line 154) “X^b is calculated as $x^b(i)-\bar{x}^b$,” instead of “X is calculated as x$(i)-\bar{x}$,”

Response: I have modified.

Comment NO.7: (line 163) “$y^b(i) - \bar{yb}$,” instead of “$y^{b(i)} - \bar{yb}$,”

Response: I have modified.

Comment NO.8: (line 165) “$(y^b)^T$” instead of “y^{b^T}”

Response: I have modified.

Comment NO.9: (line 169) “$X^a = X^b[(k-1)\widetilde{P^a}]^{1/2} = X^aW^a$” instead of “$X_a = X_b[(k-1)\widetilde{P^a}]1/2 = XaWa$”

Response: I have modified.

Comment NO.10: (line 170) “X^a” instead of “X_a”

Response: I have modified.

Comment NO.11: (line181) “according to corresponding uncertainty in MEIC inventory” Please clarify the uncertainty percentages in the MEIC inventory (e.g., $PM_{2.5}$, PM_{10}, BC, and OC at 80%, 50%, 100%, and 250%, respectively, when I roughly calculated based on the standard deviations in

the Figure S3).

Response: I have clarified uncertainty percentages in manuscript.

Comment NO.12: (line214) " $y^{b}_{PM_{2.5}}$ "instead of " $y^{f}_{PM_{2.5}}$ "

Response: I have modified.

Comment NO.13: (line 222) "$y^{b}_{PM_{10}}$ "instead of " $y^{f}_{PM_{10}}$ "

Response: I have modified.

Comment NO.14: (line 227) "$y^{b}_{PM_{10-2.5}}$ "instead of " $y^{f}_{PM_{10-2.5}}$ "

Response: I have modified.

Comment NO.15: line(261-263) italic style "…Δ*l*…*L*…*L*…" instead of "…Δl…L…L…"

Response: I have modified.

Comment NO.16: line (297) "where $RMSE^{a}$ and $RMSE^{f}$ is" instead of "where $RMSE^{f}$ and $RMSE^{a}$ is"

Response: I have modified.

Comment NO.17: line (307) "air quality index" Please add references for the air quality index and explain why AQI is used instead of direct concentrations.

Response: The AQI is a comprehensive indictor of overall air pollution and criterion for severe haze events (Zhan et al., 2018, Bao et al., 2015).

Zhan, D., Kwan, M., Zhang, W., Yu, X., Meng, B., and Liu, Q : The driving factors of air quality index in China, Journal of Cleaner Production, 197, 1, 1342-1351, https://doi.org/10.1016/j.jclepro.2018.06.108, 2018.

Bao, J., Yang, X., Zhao, Z., Wang, Z., Yu, C., and Li, X.: The spatial-temporal characteristics of air pollution in China from 2001-2014, Int. J. Environ. Res. Publ.Health, 12 (12), 15875e15887, https://doi.org/10.3390/ijerph121215029, 2015.

Comment NO.18: line (311) and Figure3(b) Please add labels for (a) and (b) and clarify axes titles for the right panel.

Response: I have modified.

Comment NO.19: (line 339 and Figure4) What the color bar values represent?

Response: The color bar represents the Gaussian Kernel Density estimation. I have modified in manuscript.

Comment NO.20: line (410-414) Since boundary conditions were used, please avoid referencing GFS accuracy for specific regions.

Response: I have avoided.

Comment NO.21: line (448-450) Simplify "also avoid the underestimation of model spread and overconfidence in the first-guess" like "avoid underestimation of model spread, which implies overconfidence in the first guess."

Response: I have modified.

Comment NO.22: line (469) "among first guess ($x^b(i)$) and analysis field ($x^a(i)$) in terms of ensemble members" instead of "among first guess and analysis field in terms of ensemble members"

Response: I have modified.

Comment NO.23: line (506 and Figure9) Does the red dots in the analysis fields (third and fourth row) represent the observation points? I would recommend just circles with a black solid line since the red color overlap with the color scale bar.

Response: Yes. I have the black dots in manuscript and supplement now.

Comment NO.24: line (554) "large spread" instead of "high dispersion"

Response: I have modified.

Comment NO.25: line (555) "background error variance" instead of "background covariance"

Response: I have modified.

Comment NO.26: line (559) "spread" instead of "divergence"

Response: I have modified.

Comment NO.27: line (560) "filter divergence" instead of "filter convergence"

Response: I have modified.

Comment NO.28: line (571) "enlarge ensemble spread" instead of "enlarge the deviations between ensemble members"

Response: I have modified.

Comment NO.29: line (575) "ensemble spread" instead of "ensemble dispersion"

Response: I have modified.

Comment NO.30: line (581) "enlarges ensemble spread" instead of "strengthens divergence"

Response: I have modified.

Anonymous Referee #2

Comment NO.1: As shown in the line 182, this study generates different emission samples by adding perturbations to the emission source inventory. This perturbation is spatially and temporally related, but this perturbation seems to be relatively strict. What is the reason for using this disturbance? Are there other disturbances that can be added to the emission inventory?

Response: Bottom-up emission inventories in compiled from activity rates and emission factors. Thus, using the same random perturbation factors, we effectively created inventory ensembles with different activity rates and emission factors. On the other hand, China has implemented series emission reduction standards in recent years, resulting in a degree of spatial and temporal correlation in emission reductions.

Comment NO.2: Picture 2 is the Flow chart of the WRF-Chem/4D-LETKF assimilation system for particulate matter. In the picture, the emission adds a disturbance to produce first-guess, and then combined with the observation and 4D-LETKF to produce the analysis result. The analysis results are used as input for the next assimilation cycle. No perturbations are added to emissions in the second cycle. Why manuscript only adds one times of perturbations into emissions at the first cycle of assimilation? Continuously adding perturbations to emissions during each assimilation cycle will increase the standard deviation in the first-guess field to produce more accurate simulation results?

Response: The research adds perturbations into emissions, actually the perturbations work in each assimilation cycle and which is enough to make most of the site information affect the first-guess field in each assimilation cycle.

Comment NO.3: In Table 3, ΔRMSE is basically negative, ΔCORR is positive, and the larger AE indicates the higher assimilation efficiency. It shows that the neighboring provinces outside the Beijing-Tianjin-Hebei region also gained greater assimilation gains. The main source of these gains is local initial field assimilation? or they come from the assimilation of the Beijing-Tianjin-Hebei region?

Response: The main source of these gains is generated from local initial field assimilation. The study conduct assimilation in two simulation domains.

Comment NO.4: In Figure8, the sensitivity experiments are conducted by the selection of ensemble member size and the length of assimilation window. What are the main reasons for the selection of 20, 40, 60 for ensemble member size and 24, 48 72 hours for the length of assimilation window?

Response: The setting of these parameters is general. Too many numbers of members will consume more computational resources. The length of the assimilation window is chosen with reference to the life cycle of the pollutant, as well as in other relevant studies.

Comment NO.5: $PM_{2.5}$ and PM_{10} have similar aerosol components, such as sea salt, and dust in GOCART scheme. And the state variables include these matters. However, As shown in Figure4 and Figure5, the RMSE of $PM_{2.5}$ in Severe-40m-48h is 31.19 and 76.83 μg m^{-3} for PM_{10}. Why the performance of PM_{10} prediction in Severe-40m-48h shows more uncertainty than those of $PM_{2.5}$?

Response: The performance of $PM_{2.5}$ in free run is better than PM_{10}. In GOCART scheme, the PM_{10} includes $PM_{2.5}$, so the uncertainty of $PM_{2.5}$ will have influence to the simulation of PM_{10}. It is still difficult to simulate the dust and sea salts. In general, there are more uncertainty in the prediction of PM_{10}.

Comment NO.6: Write the formulation about correlation coefficient and root mean square error.

Response: I have modified.

Comment NO.7: Replace the “peak” in line 308 by the “peak value of”.

Response: I have modified.

Comment NO.8: Add the units in Figure 5.

Response: I have modified.

Comment NO.9: Replace “air quality index” in line 307, 319, 611 by “AQI”.

Response: I have modified.

Comment NO.10: Add the units in Figure6.

Response: I have modified.

Comment NO.11: Add the units in table 3.

Response: I have modified.

Comment NO.12: Replace “larger” in line 314 by “all larger”, replace “below” in line 315 by “all below”.

Response: I have modified.

Comment NO.13: Replace “adjustment” in line 324 and 600 by “initial condition adjustment”.

Response: I have modified.

Comment NO.14: Replace “Beijing-Tianjin Hebei” in line 408 by “BTH”.

Response: I have modified.

Comment NO.15: Replace “provide” in line 105 by “providing”.

Response: I have modified.

Comment NO.16: Replace “event” in line 106 by “events”.

Response: I have modified.

Comment NO.17: Replace “dimensions” in line 426 by “dimensions”.

Response: I have modified.

Comment NO.18: Replace “investigated” in line 628 by “the investigated”.

Response: I have modified.